NOTE

This Microfiche meets the quality requirements of the American National Standard for Microfiche of Documents (ANSI PH5.9-1975). However, because of imperfections in the document from which it was made, it contains images that are illegible or difficult to read. The source document for this microfiche is the best copy that could be obtained prior to filming.

CIS Congressional Information Service Washington, D.C.

The document was prepared under a contract with the Office of Technology Assessment, Congress of the United States, as an input to the assessment of distance learning. The views expressed by the authors are not necessarily those of OTA, the Board, the OTA Advisory Council or individual members there of.

EVALUATING DISTANCE LEARNING TECHNOLOGY

Richard E. Clark
University of Southern California

May 1989

Prepared for the Office of
Technology Assessment, U.S. Congress
Contract #L3-2435.0

J952-52

EVALUATING DISTANCE LEARNING TECHNOLOGY

Richard E. Clark
University of Southern California

Final Draft
May 23, 1989

The history of distance learning technology claims (1, 2, 3) and counter claims (4, 5, 6) have led many to accept the need for a change in the way we evaluate new technologies for distance learning. The number of new and complex technological devices that could be applied to distance learning is increasing. Stanford professor Larry Cuban, in his recent book Teachers and Machines (2), cautions that "determining what levels of [technology] use now exist is like trying to snap a photograph of a speeding bicyclist".

With adequate evaluation in place, we may be able to "tune" existing technologies so that they meet our needs, anticipate new developments, and settle disputes, in time to plan and operate rational, cost-effective K-12 local, regional and national distance learning systems. Underlying all evaluation plans are beliefs about how we employ technology so that we enhance the delivery of instruction and the quality of learning experiences. At the heart of every evaluation plan is a curiosity about how new technologies can increase student access to quality instruction and thereby increase their academic achievement, motivation and value for learning.

The purpose of this discussion is to a) discourage evaluation questions which have not proved useful in the past, b) suggest that future evaluations distinguish between the effects of delivery and instructional technologies, c) offer some generic evaluation plans, questions and examples associated with delivery and instruction; and d) discuss issues related to the evaluation of distance learning cost-effectiveness.

FORMING AND ASKING EVALUATION QUESTIONS

Evaluation is the process by which we judge the "worthwhileness" of something. Since our values govern all evaluation activities, we need to be clear about the kinds of distance learning evaluation questions that will meet the needs of our schools and communities. The questions we decide to ask about distance learning and the evaluation instruments we employ will necessarily keep us ignorant about some matters while informing us about others. Evaluation questions carry implicit assumptions and beliefs about the significance of different elements of distance learning and their impact on desired outcomes. For example, if we ask whether a new teaching medium produces more student achievement than traditional media, we have assumed that media <u>are able to</u> influence student achievement -- an assumption which has been seriously questioned (4,5,6).

One of the most important recommendations underlying this discussion is that all evaluations should explicitly investigate

the relative benefit of two different but compatible types of distance learning technologies found in every distance learning program. One technology influences the delivery of schooling and another technology influences instruction. These two technologies are typically confused in most distance learning evaluations. Technology benefits caused by instructional technology are attributed to delivery technologies, and vice versa. The confusion of technological benefits can lead to inappropriate policy decisions. At the root of the confusion one finds different definitions of "technology".

Which Technology for What Purpose?

It its most general sense, The term "technology" suggests the application of science and experience to solving problems (8). The major obstacle in our past struggle to understand the contribution of new technology in distance learning is that we have confused the contributions of these two different technologies.

One distinct class of technologies results from the application of various scientific and engineering principles to hardware that records and transmits instruction. These "media" technologies are associated with the physical sciences that have produced the new electronic media (e.g. fiber optics, television, computers). Delivery technologies increase student and teacher access to learning resources which is one of the most important goal of distance learning.

A second type of technology applies various social science principles to suggest teaching methods and curriculum choices. This instructional technology draws primarily on research in teaching, learning and motivation to enhance student achievement. The "products" of an instructional technology are new instructional design theories (9), teaching methods and motivational strategies (4) which can be embedded in "courseware" (instructional materials) for distance learning. One purpose of this discussion is to recommend that all evaluations of distance learning programs attempt to provide reliable and valid determinations of the separate influence of delivery and instructional technologies.

Separating Delivery and Instructional Questions

Support for a separate consideration of delivery and instructional technologies in evaluation is well established in the research literature but rare in evaluations or program planning. Wilbur Schramm, the most established reviewer of media studies in education, concluded (10) that "...learning seems to be affected more by what is delivered than by the delivery medium" (p.273). For the past two decades at least, most of the exhaustive analyses of research that compared the learning benefits of different media (4, 5, 6, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 22, 23) could be summarized with the analogy that media "... do not influence learning any more than the truck that delivers groceries influences the nutrition of a community" (11, p.3). Distance

learning media are vehicles that transport instruction to students. The choice of vehicle influences the important outcomes of student access, and the speed or cost of the delivery but not the learning impact of the instruction that is delivered to the "consumer". Delivery vehicles indiscriminately carry helpful, hurtful and neutral instruction.

CHOOSING CRITICAL INDICATORS FOR DISTANCE LEARNING TECHNOLOGIES

Among the specific issues that must be addressed in future evaluations of local applications of distance learning technology are: "What aspects of evaluation planning enhance the usefulness of information for decision makers?", and "How might we collect information which will aid in our judgement about the different influences of the delivery and instructional features of the program?" (12, 21, 24). While a number of evaluation concerns apply to some but not all programs, three generalizations seem useful to all, 1) Adopt an early concern for evaluation, 2) Use a multi-level evaluation plan, and 3) Conduct formal cost-effectiveness analyses.

Adopt An Early Concern For Evaluation

Both evaluation specialists and administrative decision makers need to be involved early and actively in distance learning system design. Past experience suggests that waiting until a system is

designed before thinking about evaluation has been common but very wasteful. It is critical to have early information about, for example, exactly what set of conditions are being replaced by new distance learning programs. One way to accomplish this would be to spend an ample amount of time during the program planning stages to carefully describe the specific problems we wish the new approach to solve. We should describe how we will measure the current conditions (e.g., a base-line measure of the existing situation --including the views and impressions of the "stakeholders") and thoroughly discuss what we believe to be the alternative solutions to the problem(s).

If an evaluation plan is developed as the program is planned and implemented a number of advantages are realized. In the area of computer assisted learning, Henry Levin (25, 26, 27) describes eight exemplary cost-effectiveness evaluation programs. Each of these good examples collected baseline measures of the problems they were trying to solve. Each of the eight programs began their concern with evaluation at the start of their planning.

Early evaluation makes it possible to determine which aspects of a distance learning program were positive and which were negative. Negative aspects can then be modified and the positive accentuated to achieve maximum benefit. For example, most distance learning programs attempt to bring a much richer set of curriculum choices and quality teaching to K-12 school programs. Program planners might begin with an analysis of options (the variety of media available to deliver new curricula) and their audience (e.g.

measure the number of "students" who would enroll in new courses). Early concern with evaluation results in the collection of information on both the need and the audience for distance learning as well as the existing alternatives. Another advantage of early evaluation involvement is that an ongoing evaluation plan can be developed. Too often programs are developed, implemented and at some later stage the program planners remember that "we have to do some evaluation". As a result, very little is learned about the program being studied which is useful either for the immediate program or for other distance learning ventures. It is early evaluation planning most often yields a useful information and yet it is a rare phenomenon. Levin noted that his search for adequate cost effectiveness evaluations was difficult. He found that only one in six published reports were adequately designed. Since most reports are not published, one suspects that early evaluation is not our typical procedure at the moment.

The second general direction that is useful for all distance learning programs is to adopt a multi-level evaluation plan.

Use a Multi-level Evaluation Plan

The two levels of evaluation that most often seem to give useful development information are to measure: 1) participant reactions, and 2) the achievement of program objectives.

1) Participant reactions to distance learning program

effectiveness is the most common (and unfortunately, often the only) level of evaluation attempted by K-12 school districts. Typically, this level of evaluation employs printed forms containing a combination of questions designed to inquire about the "feelings" and "impressions" of different groups who are involved in the program. A common question is "How would you rate the quality of the teaching in this program?" (typically rated on a five point scale that ranges from Exceptional through Average to Poor). Items such as "List what you think are the STRONG [or WEAK] points of the program?" permit the respondent to write in personal views and comments. Questionnaire forms are most often used for reactions because they protect the anonymity of respondents and therefore, we presume, increase the candor of the responses. Forms are often sent to all of the program participants but are filled out and returned by only a small percent of those who receive them.

Participant Reactions are useful provided that they do not serve as they <u>only</u> level of evaluation data. They should be used primarily to uncover both informal participant impressions and <u>unanticipated</u> benefits and problems. Reaction items should be divided between those that deal with the medium (e.g., ease of access, reliability or technical quality of transmission or machines, space allocation issues) and those associated with the instruction (e.g., the quality of teaching, how things learned in the program were used outside of class).

<u>Unanticipated results</u>. The advantage of collecting participant reactions is that program mangers get informal impressions of the

programs and often uncover unanticipated results. For example, the Northeastern Utah Telelearning project, which uses microcomputer audiographic instruction transmitted between remote schools over telephone lines, found an unexpected problem because they used open-ended reaction forms. Students complained that in the early stages of the program it was very difficult to contact a teacher to get help when it was needed. On the other hand, the Interact Instructional Television Network in Houston Texas used a similar instrument and discovered an unexpected positive outcome of their project. It was observed that students in small television reception rooms tended to help each other a great deal during the instructional program. They could help while the program was continuing without disrupting the teacher or other students. This "peer tutoring" seemed to be having a positive impact on student learning and motivation. Upon closer inspection the tutoring activity seemed to be due to the fact that the microphone which the students used to communicate with the teacher at another location had to be turned on to function. When the microphone was turned off, the students could consult among themselves before they turned it on to answer a question or discuss a point with the teacher. When following up on the peer tutoring finding, the Houston project uncovered the fact that some of the tutors hired to supervise the student television reception rooms were demanding that the students "keep quiet" which discouraged the peer tutoring. The tutors had assumed that talking indicated a "discipline problem" and had to be corrected by their supervisors. Once

discovered, the peer tutoring can be encouraged and its barriers can be eliminated (e.g., though adjusting the training of tutors in the Houston example).

Reliability and Validity of Questionnaires. The disadvantages of participant reaction data is that it is seldom gathered in such a way that it can be considered either a reliable or valid reflection of the program. This problem is not serious. Unreliable information can still provide useful information as it did in the Houston project. However, questionnaire data can be representative if evaluators select a random sample of participants large enough to engage a meaningful number of each group involved in the project. To increase participation evaluators have found it useful to send each randomly chosen participant a card telling them that they have been selected, that their response is vital and to expect the questionnaire soon. Follow up notes to all those chosen can encourage laggards to send in their forms without violating anonymity. Depending on the numbers involved in the entire program, a small (five to ten percent) random sample of participants can give a very accurate impression of the reactions of the entire group.

Questionnaires should be used at various stages in the program development, including very early on. Unanticipated problems and benefits uncovered by questionnaires usually require much more careful study. For example, when students in most distance learning programs are asked, a majority will typically state that

they would not continue to elect a distance learning option if they could choose a "traditional class" as they did in the Northeastern Utah Telelearning Project. The fact that students would elect a traditional program if one was offered does not indicate that the Utah project failed. Upon closer inspection it is often found, as is suspected in the Utah case, that students sometimes feel isolated in distance learning settings and would therefore select more traditional options for social and "nonacademic" reasons. This is particularity true of middle and high school students. They typically have strong social needs which are not always met in distance learning programs.

Other problems which can be spotted using the "early warning system" of reaction questionnaires are communication problems between participants, the extent and impact of technical difficulties, inappropriate implementation of plans and opportunities to extend the program into new areas. Yet in even the best of circumstances, reaction forms will not give solid information about the achievement of most delivery and instructional goals. For this purpose, programs need to adopt a second level of evaluation.

2) Achievement of Program Objectives is the second and most substantive evaluation goal. Formal measurement of objectives is usually considered by evaluation specialists to be the most crucial information to be gathered. Objectives should be divided into at least two categories, those associated with delivery and those

associated with instruction. One category of outcome which is common to both types of technology is cost-benefit. The discussion turns next to outcomes specific to instructional technology, then to delivery technology outcomes and finally to cost-effectiveness measurement.

Instructional Technology Objectives include student learning, motivation, transfer of knowledge and values. These important goals are influenced by the "courseware" or instructional programs that are developed and/or chosen and transmitted to distant learners. In most cases, instruction is designed by teachers. In some cases, already developed courseware is purchased and transmitted to remote sites. The instructional decisions that are embedded in each lesson influence student learning and motivation. Different teaching method and curriculum options have very different effects on student learning which might be explored in evaluation. So, distance learning evaluation might include at least the following four types of questions related to instructional technology.

I. "Which of the curriculum and teaching method choices in a given distance learning program impacted student achievement and subsequent ability to use (transfer) the knowledge acquired outside of the instructional setting?"

Achievement can be tested with teacher or publisher achievement tests. Increasingly schools are interested in the

extent to which students transfer what they learn outside of school. Transfer might be estimated by open ended questions on reaction forms. If the school district has other schools receiving similar curricula from different delivery forms, an obvious opportunity exists to check on any achievement or motivation differences between the options. When possible, alternative teaching methods and curriculum choices should be explored in order to maximize the learning of different kinds of students. For example, highly structured and supportive instruction might be contrasted with a more "learner directed" and discovery approach to curriculum (12). Many programs have found that students who are anxious or have learning problems profit a great deal from added structure and support. Whereas students who are more independent and able tend to benefit more from a discovery approach (11).

II. "What impacted student and teacher motivation to learn and invest effort in making this program a success?"

Current theories of motivation have introduced a very novel element in distance learning programs and evaluation. Formerly, it was thought that media choices greatly influenced both student and teacher motivation. Now it is understood that motivation is influenced by beliefs and expectations and is therefore due to "individual differences in beliefs about media" and not to the media per se (5, 11, 22). Yet, it is a common belief that students are excited and teachers are threatened by new media (2). There

is recent but solid evidence that when students expect that a new medium will make learning easier and more "entertaining", they like it. However, there is good evidence that their liking does not lead them to work harder (3, 5). Quite to the contrary, the more they think a medium makes learning "easy" the less effort they will invest to learn (18, 22). This effect has been explained as a misjudgment about the kind of effort that is required to learn based on our previous experience and expectations. For example, American students typically assume that television is an "easier" medium than books or teachers, probably because of their use of the medium for entertainment. This reaction on the part of our students is quite different than that of Israeli students who, on the average, have been found to invest more effort in television because their early experiences with television have been less entertaining and more demanding intellectually (22).

There is additional evidence that students will not invest effort if they believe a medium to be very difficult. With American children, this is sometimes the reason for their lack of willingness to read (18, 22). So the greatest motivation is invested in media and instructional programs that are perceived as being moderately difficult. This evidence would suggest that one way to influence student motivation would be to select "moderately difficult" media. However the evidence also suggests that student and teacher beliefs about media difficulty change over time, sometimes radically (5). The more stable predictor of motivation seems to be student beliefs about their own ability and the demands

placed on them by different instructional tasks (18). This would suggest that we should evaluate the students perceptions and beliefs about the learning tasks contained within the media employed by distance learning programs and their own self efficacy as learners. This form of evaluation could be embedded in reaction questionnaires.

> III. "Which of the curriculum and teaching method choices in a given distance learning program impacted student and teacher values for what was learned and subsequent motivation to teach and learn and to use what was learned outside of the instructional setting?"

Reaction questionnaires which are carefully constructed and administered will give a good indication of student and teacher values related to the program, teaching and the curriculum. Negative value statements do not always reflect negatively on the program (recall the students in the Utah project who liked traditional classrooms better than distance learning because of social opportunities). Generally one hopes to foster a positive value for learning and new curriculum options with distance learning. Shifts in attitude that result from changes in the program can be monitored if reaction forms are sent periodically (every few months) throughout the development stages.

IV. "Which of the curriculum and teaching method choices in a given distance learning program impacted the cultivation of different kinds of knowledge including procedural skills and higher order thinking, learning-to-learn and metacognitive skills?"

While higher order skill learning is more difficult to assess than ordinary "achievement", some programs have been successful in this area. Perhaps the most exciting current example in wide use is to be found in the HOTS program developed by Dr. Stan Pogrow at the University of Arizona (24). HOTS (Higher Order Thinking Skills) is a successful and widely disseminated program for Title I students. Teachers in the program use computer lessons, class exercises and discussion to increase the thinking and study skills of students. Evaluation involves the ongoing use of standardized tests, noting changes in the quality of questions students ask and analyses of their class assignments. While a few formal measures of thinking and study skills exist (and more are being developed), program managers might consult with evaluation specialists about selecting and developing tests to measure problem solving and study skill development (24,25).

While learning, values and study skills are important instructional outcomes for distance learning, the delivery technology will influence yet another type of outcome.

Delivery Technology transmits various forms of instruction to students. The recent introduction of computers to schools has resulted in more attention to technology delivery benefits (15, 24). Evaluation questions associated with delivery technologies include attempts to assess the effect of medium on 1) student access to a greater variety of curriculum choices, 2) school or program utilization of resources, and 3) the reliability of delivery choices. Questions one typically finds in the evaluation of media include:

I. "Did the distance learning media maximize student access to new, and/or high quality courses and teaching when compared with other choices?"

Access to new or beneficial courses and instructional techniques or teachers is one of the primary objectives of most distant learning programs. Collecting access data often involves comparisons between different ways to deliver courses or the size of enrollments in classes both before and during the implementation of the program. For example, the Share-Ed program in Beaver County Oklahoma used a new fiber optic network to provide new curriculum to rural schools. They collected participant reactions on the advantages of the increased curriculum choices offered to students who allowed to take college credit courses in high school as a result of the new system. These reactions, when combined with baseline and process data on actual enrollments, provide good

evidence of the extent of access provided by the innovation. Evaluators should carefully consider increased or enhanced access of minority, older or widely dispersed student groups.

While "access" usually suggests the availability of new curriculum options, it can also imply teacher access to students on a more personal level. Teachers in the Houston Texas InterAct Instructional Television systems report problems with their personal and immediate access to students during instruction in order to "check their reactions or mood" and adjust their teaching accordingly. Whereas teachers using computer delivered courses often report increased "individualized" access to students and enjoy the opportunity to "watch them learn".

> II. "Did the media influence the <u>utilization</u> of school and community educational resources (e.g.space, equipment, skilled teachers, new courseware developed at one site but not readily available at others)?"

It is often the case that because distance learning programs are recorded and distributed to many different sites, the best teachers are made available to many more students. Evaluators might track statistics about how the background and/or training of teachers in distance programs compare with district averages. An instance of a different kind of utilization is to be found in the Beaver County Oklahoma Share-Ed program. The local telephone company was installing fiber optic communication lines to improve

local service. The system was capable of handling far greater transmission volume than the existing usage anticipated in the communities served. The school system's use of fiber optic lines for television and voice transmission for distance learning utilized unused space on the system. Since distance learning courses are often provided to fewer students per school than the average course, they often make use of under utilized rooms (e.g. storage spaces) and equipment.

III. "Are distance learning media more reliable than other alternatives?"

One of the primary concerns expressed by the critics of distance media is their technical reliability. In the Beaver County Oklahoma television system for example, the reaction forms used in evaluation only picked up technical problems when the students were asked to describe "weak points" of the system. None of the administrators noticed technical problems, eleven percent of the teachers mentioned reliability, but thirty six percent of the students responded to the reaction form by going into detail about microphone feedback, distracting equipment, out-of-focus pictures, equipment noises and color problems. This difference in reporting reliability problems probably stems from the amount of experience each group had with the actual television transmission. However, program evaluation should establish regular checks by technical staff on these problems in order to judge the severity

of participant reactions and make repairs when necessary. When technical transmission problems are not solved, they can decrease achievement scores and reduce participant commitment to the system.

In all successful distance learning programs, delivery (media) technology and instructional technology must work together. The delivery features of new media must be employed so that they will eventually save precious educational resources. Curriculum and instructional design must be utilized so that they support the effective learning and transfer of important concepts. Instruction must be developed to reflect the special delivery characteristics of different media. In addition however, communities and funding agencies are increasingly concerned not only with the effectiveness but also the cost of distance learning programs. Cost is a "goal" or "outcome" of both delivery and instructional technologies.

COST-EFFECTIVENESS EVALUATION

During an evaluation of the separate delivery and instructional value of distance learning program effectiveness, cost data should also be collected. This parallel activity allows us to combine "effectiveness" (i.e., delivery and instructional outcomes) with "cost" data to provide cost-effectiveness information to decision makers.

In many ways, cost-effectiveness ratios are the most interesting information we can supply to school officials, taxpayers and their elected representatives. Limited educational

resources will eventually require a much greater emphasis on both the monetary and time cost of new programs.

DELIVERY TECHNOLOGY COST

Evaluations that precede the introduction of new media should explore the costs of various alternatives. In many cases, older technologies (e.g., tutors, books, cassette television programs, the mail system) are cheaper in monetary cost but very "expensive" in delivery time and reliability. Evaluations of costs should always consider trade-offs with cheaper and more traditional delivery options. There is evaluation data which indicates, for example, that tutors who are trained and paid minimum wage are much cheaper than computers for some instructional purposes (25).

Evaluations which are conducted during the introduction and maintenance of a distance learning program are advised to adopt the "ingredients" costing approach described below.

INSTRUCTIONAL TECHNOLOGY COST

There are a great variety of different school and community goals that influence evaluation criteria under the general heading of instructional effectiveness costs. The cost involved in increasing student motivation, learning and transfer are being questioned with greater frequency. School districts may wish to consider collecting cost data which will aid policy makers. The

development of an instructional technology yields a variety of teaching, motivation and transfer outcomes at very different monetary costs.

Besides monetary cost, schools are increasingly interested in the time costs associated with the mastery of different learning or performance goals. Some types of learning tasks consume much more "teaching time" and/or "learning time"(5). For example, it takes much longer to teach a student study skills than to teach memorization of facts. It also takes longer for a student to learn procedural knowledge to the point where it becomes automatic -- about 100 hours of practice for even simple procedures is the current estimate. Therefore, there will be more and more emphasis on the time costs of different instructional technology options. In many areas, the cheapest option is not necessarily the best. In the same way, the quickest option among instructional technologies is not always the best. Students who learn faster do not necessarily learn better. The new "cognitive" learning theories provide the insight that it may be more important to know <u>how</u> students reach learning goals than to know that they get correct answers on examinations. It often takes longer for students to learn in such a way that their correct answer on a test reflects "deep cognitive processing" and the exercise of "higher order cognitive learning skills", than to take a surface level shortcut. Educators need to be wary of focusing evaluations on time savings at the expense of the quality of learning.

Generally, once a distance learning team has worked out the list of goals associated with both monetary and time costs, an evaluation design can be chosen. One of the first issues to be confronted is the choice of how the data reflecting costs will be gathered. While there are a number of methods, one seems particularly applicable to both delivery and instructional technologies - Levin's (25, 26, 27) ingredients method.

THE INGREDIENTS METHOD OF DETERMINING COSTS

While there are a number of emerging ways to determine local costs and efficiencies, one of the soundest and most comprehensive is the "ingredients method" developed by Henry Levin at Stanford University (25, 27). It " requires identification of all of the ingredients required for the ... [distance learning] intervention, a valuation or costing of those ingredients and a summation of the costs to determine the cost of the intervention" (27, p. 3). In the K-12 setting, cost is defined as the value of what is given up by using resources in one way rather than for its next best alternative use. For example, if teacher time is given up then it may not be used for other purposes. Therefore, the cost of teacher time is assessed by assigning a value to what is lost when teachers are assigned to distance learning technology programs.

The ingredients method is implemented in two stages. In the first stage, all necessary program ingredients are listed. The identification of ingredients requires that we list distance

learning program necessities associated with five categories: 1. personnel, 2. facilities, 3 equipment, 4. materials and supplies and 5. all other. In the second stage, each of the ingredients listed in each of the five categories is valued.

Space limitations preclude a complete description of the ingredients method but a review of Levin (26,27) will provide most of the information needed to determine ingredient costs. Levin gives specific technology examples which are very relevant to the kinds of programs now evolving in many schools and he urges complete listings of ingredients. For example, he requires that all "donated" time of volunteers and outside organizations be included as a personnel ingredient if it is necessary for the conduct of the program. He reasons that failure to cost donated time will give an unrealistic picture of the "replication" expense. He also claims (26) that, in the rare instance where one finds a complete costing of technology-based programs, one often finds evidence that the organizational climate greatly influences cost-benefit ratios. Some organizational plans seem to be much more efficient than others.

CONCLUSION

In the past, distance learning evaluations have typically been conducted as "afterthoughts" and have relied heavily on reaction questionnaires which are unreliable and nonrepresentative of the participants involved. Even when evaluations attempted to collect

information about changes in student achievement, questions were asked which confused the separate contributions of delivery media and instructional technology.

In order to identify the strong features of distance learning programs and eliminate weak features, more robust evaluation plans must be adopted in the future. Three features are recommended: First, evaluation should begin at the start of distance learning program planning. An early commitment to evaluation will provide much more useful information about the strengths of a program as it develops. Changes can be made during the formative stage in time to strengthen the plan. The second recommendation is that all programs should adopt a multi-level evaluation plan. The different role of qualitative (e.g. questionnaire) and quantitative (e.g. student achievement scores, monetary costs) data should be decided. Delivery and Instructional evaluation should be separated and a variety of goals assessed. Finally, new techniques are available for cost-effectiveness evaluation of distance learning programs. Levin's "ingredients" method is suggested.

REFERENCES

1. Benjamin, L.T., (1988) A history of teaching machines, American Psychologist, 43(9), 703-712.

2. Cuban, L. (1986) Teachers and machines: The classroom use of technology since 1920. New York: Teachers College Press.

3. Saettler, P. A., (1968), A history of instructional technology. New York: McGraw Hill.

4. Clark, R. E. & Salomon, G. (1986) Media in teaching. In Wittrock, M. C. (Ed.), Handbook of Research on Teaching. Third edition. New York:Macmillan.

5. Clark, R. E. & Sugrue, B. M. (1988) Research on instructional media: 1978-1988. In D. Ely (Ed.) Educational Media and Technology Yearbook volume 14. Denver, CO:Libraries Unlimited.

6. Gardner, H. & Salomon, G (January, 1986) The computer as educator: Lessons from television research. Educational Researcher, 13-19.

7. Clark, R. E. (1989), Current progress and future directions in research on instructional technology. Educational Technology Research and Development. 37(1), 57-66.

8. Heinich, R. (1984) The proper study of instructional technology. Educational Communication and Technology Journal, 32(2), 67-87.

9. Reigeluth, C. (1983) Instructional design: Theories and models. Hillsdale N.J.: Lawrence Earlbaum Associates.

10. Schramm, W. (1977). Big media, little media. Beverly Hills, CA:Sage Publications.

11. Clark, R. E. (1983) Reconsidering research on learning from media. Review of Educational Research, 53(4), 445-459.

12. Clark, R. E. (1985) Evidence for confounding in computer-based instruction studies:Analyzing the meta-analyses. Educational Communication and Technology Journal, 33(4), 249-262.

13. Allen, W. H. (1971) Instructional media research: Past, present and future. A V Communication Review, 19, 5-18.

14. Jamison, D., Suppes, P., and Wells, S. (1974) The effectiveness of alternative instructional media: A survey. Review of Educational Research, 44, 1-68.

15. Kearsley, G., Hunter, B., and Sidel, R. J., (1983) Two decades of computer based instruction:What have we learned? T.H.E. Journal, 10, 88-96.

16. Levie, W. H. and Dickie, K. (1973) The analysis and application of media. In R. M. W. Travers (Ed.) Second handbook of research on teaching. Chicago: Rand McNally.

17. Lumsdaine, A. (1963). Instruments and media of instruction. In N. L. Gage (Ed.) Handbook of research on teaching. Chicago:Rand McNally.

18. Mielke, K. (1968). Questioning the questions of ETV research. Educational Broadcasting Review. 2, 6-15.

19. Salomon, G. (1981) Communication and Education. Beverly Hills CA: Sage Publications.

20. Winn, W. (1982). Visualization in learning and instruction. Educational Communication and Technology Journal, 30(1), 3-25.

21. Baker, E. L. (1989) Technology assessment: Policy and methodological issues. Air Force Systems Command, "Proceedings of the 2nd intelligent tutoring systems research forum", 151-158.

22. Salomon, G. (1984) Television is "easy" and print is "tough": The differential investment of mental effort in learning as a function of perceptions and attributions. Journal of Educational Psychology. 76(4), 647-658..

23. Clark, R. E. & Voogel, A. (1985) Transfer of training for instructional design. Educational Communications and Technology Journal, 33(2), 113-123.

24. Congress of the United States, Office of Technology Assessment, Power On: New tools for teaching and learning. (Washington, D.C., U.S. Government Printing Office, September 1988).

25. Levin, H.H. (1983) Cost Effectiveness: A primer. Beverly Hills, CA: SAGE Publications.

26. Levin, H. M. and Meister, G. R. (1985) Educational technology and computers:Promises, promises, always promises. Report No. 85-A13. Stanford, CA:Center for Educational Research at Stanford. School of Education, Stanford University.

27. Levin, H.H. (May 1988) The economics of computer-assisted instruction. In press for a special issue of the Peabody Journal of Education to be edited by J.J. Bosco, Tate Center, Western Michigan University.